

Verplichte datadeling voor specifieke sectoren noodzaak

In sommige sectoren kan verplichte datadeling voorkomen dat datagedreven markten omslaan naar een monopolie. Lopende beleidsinitiatieven laten zien hoe de publieke belangen van concurrentie, innovatie en gegevensbescherming met elkaar verenigd kunnen worden.

INGE GRAEF

Universitair docent rechten aan Tilburg Law and Economics Center (TILEC), Tilburg University

JENS PRÜFER

Universitair hoofd-docent economie aan TILEC, Tilburg University

In beleidskringen wint het inzicht terrein dat data een belangrijke motor van groei zijn. Zo heeft de Europese Commissie zich, door middel van het *European data economy*-initiatief, als doel gesteld het delen en hergebruiken van data te bevorderen, om zo innovatie te stimuleren (EC, 2017). En ook het Nederlandse Ministerie van Economische Zaken en Klimaat erkent, in het nieuwjaarsartikel in *ESB*, dat data die meer wijdverbreid beschikbaar zijn, een belangrijke bijdrage aan de samenleving kunnen leveren.

Daarmee neemt het ministerie de visie over op datadeling die Prüfer en Schottmüller in het artikel 'Competing with big data' (Prüfer en Schottmüller, 2017, hierna: PS) uiteenzetten: "Door geanonimiseerde data over klikgedrag breder beschikbaar te stellen, kunnen ook andere partijen in andere markten met behulp van deze data verder innoveren, en kan men de concentratie van grote internetbedrijven in deze andere markten voorkomen." "Het gaat dan bijvoorbeeld om de markt voor digitale kaarten, retail en in de toekomst ook om zelfrijdende auto's." (Camps, 2018). Tegen de achtergrond van de toenemende steun onder beleidsmakers voor datadeling, zetten wij in dit artikel allereerst de belangrijkste argumenten uit PS uiteen om aan te tonen waarom en wanneer het delen van informatie essentieel is. Vervolgens bespreken we de beleidsmogelijkheden voor het opleggen van verplichte datadeling.

CONCURRENEN MET BIG DATA: EEN TERUGBLIK

PS introduceerden *datagedreven netwerkeffecten* als een nieuw economisch mechanisme, definieerden *datagedreven markten*, en verduidelijkten daarmee welk type gegevens – namelijk *gebruikersinformatie* – de 'olie van de data-economie' is (The Economist, 2017).

Gebruikersinformatie zijn gegevens over de voorkeuren of kenmerken van de gebruikers van een dienst, zoals bijvoorbeeld de geografische locatie van een gebruiker van een mobiele app of de voorkeur van een zoekmachinegebruiker die op zoek is naar een Siciliaans restaurant in Amsterdam. Deze gegevens zijn privé-informatie van het bedrijf dat de dienst levert en dat de informatie verzameld heeft. Die informatie over de gebruiker kan het bedrijf vervolgens inzetten als innovatie-input, om de dienst beter af te stemmen op de voorkeuren van de gebruiker. Hierdoor neemt de kwaliteit van de dienst in de toekomst toe.

Een hogere initiële vraag naar een dienst vermindert dus de marginale kosten van innovatie: het is dan goedkoper om één extra eenheid product- of dienstkwaliteit te produceren, zoals ervaren door gebruikers. Dit economische mechanisme wordt een *datagedreven indirect netwerkeffect* genoemd en markten die onderhevig zijn aan deze effecten worden *datagedreven markten* genoemd. Cruciaal hierbij is dat het verzamelen van gebruikersinformatie een bijproduct is van de zoekopdracht en transactie door de gebruiker en vrijwel gratis is voor het bedrijf dat de dienst levert. Dit creëert een toetredingsbarrière voor potentiële concurrenten.

Markten waarin het verzamelen van gegevens over gebruikers positieve marginale kosten kent of waar de data geen informatie over de voorkeuren of kenmerken van gebruikers omvatten, zijn daarentegen niet-datagedreven. In deze markten kunnen gegevens ook belangrijk zijn, maar is de invloed van data gebaseerd op gangbare economische mechanismen en is de kans op monopolievorming meestal minder groot.

De belangrijkste conclusie van PS is dat datagedreven markten bijna altijd kantelen, en zich richting een monopolie bewegen. Cruciaal is dat innovatieprikkels in een gekantelde markt daarna laag zijn. Nieuwe toetreders moeten concurreren met een gevestigde speler die lage marginale kosten van innovatie heeft, vanwege de grote voorraad aan gebruikersinformatie. Bovendien kent het dominante bedrijf de angst van potentiële uitdagers vanwege zijn datavoordel. Vandaar dat na het kantelen van de markt niet alleen de concurrentie *op* de markt, maar ook de concurrentie *om* de markt sterk is verminderd.

Deze theoretische uitkomst wordt ook empirisch ondersteund (He et al., 2017). Preston McAfee, de hoofdeconoom van Microsoft, stelde onlangs op basis van zijn empirisch werk naar de kwaliteit van zoekmachines: “Zelfs op de schaal van het internet maken meer gegevens het resultaat van zoekopdrachten beter” (McAfee et al., 2015). Junqué de Fortuny et al. (2017) bevestigen dat data onderhevig zijn aan een afnemend rendement, maar een strikt positieve invloed hebben op de nauwkeurigheid van voorspellingen. Hoewel de toegevoegde waarde van meer gegevens over gebruikersvoorkeuren en -kenmerken dus kleiner wordt naarmate meer gegevens verzameld zijn, is de waarde van data dus altijd positief en van invloed op de kwaliteit van het product en daardoor op de marktuitskomsten.

Een ander concept dat PS introduceerden, is het idee van *verbonden markten*, om zo de bedrijfsstrategieën van de meest succesvolle internationale bedrijven te begrijpen. Datagedreven bedrijven kunnen verschillende markten met elkaar verbinden als de gebruikersinformatie die zij op de ene markt hebben verzameld ook waardevol is in een andere markt. Sommige zoekopdrachten van zoekmachines hebben bijvoorbeeld betrekking op geografische informatie, dat ook waardevol is bij het aanbieden van een gepersonaliseerde kaartendienst. PS toonden aan dat een bedrijf met een datagedreven bedrijfsmodel op die manier op de lange termijn elke traditionele markt kan domineren als de kosten van toetreding niet te hoog zijn. Gebruikersinformatie speelt een belangrijke rol in dit proces dat zich herhaaldelijk kan voordoen, waardoor een domino-effect ontstaat over verschillende markten heen.

INDIRECTE NETWERKEFFECTEN

Critici zullen beweren dat we dit al eerder gezien hebben en dat datagedreven indirecte netwerkeffecten hetzelfde zijn als leercurve-effecten. Vanuit die gedachte voorspelden pessimisten tien jaar geleden immers ten onrechte de ‘eeuwige’ dominantie van Microsoft. Het verschil tussen die twee effecten is dat wanneer er sprake is van een leercurve-effect, en als twee concurrenten ongeveer dezelfde achtergrond hebben (wat zich vertaalt in dezelfde output in het verleden), beide bedrijven ook ongeveer gelijke productiekosten hebben. Het gevolg is dat zij de markt voor altijd zullen delen. Bovendien is een leercurve meestal een tijdelijk voordeel. Zodra een bedrijf de technologische bovengrens bereikt, zal zijn leercurvevoordeel gestaag afnemen en kunnen toetreders of kleinere bedrijven de achterstand inlopen.

In een datagedreven markt richt een bedrijf zich echter op investeringen in kwaliteit in plaats van op de prijs waardoor de verschillen tussen bedrijven permanent zijn vanaf het moment dat de markt gekanteld is en beide bedrijven het optimaal vinden om niet meer te investeren. Een nieuwkomer of een kleinere speler kan dan niet winstgevend zijn en dezelfde kwaliteitsinvestering doen die nodig is om de dominante speler te evenaren.

De belangrijkste nieuwe bevinding vanuit het oogpunt van het mededingingsbeleid is dat in een markt waar sprake is van een leercurve-effect, bijvoorbeeld in de vliegtuigbouw, een kleiner bedrijf B de belangrijkste technici van dominant bedrijf A kan overnemen – en daardoor

relevante kennis opdoet – waardoor het vervolgens fel kan concurreren met bedrijf A. In een datagedreven markt is het zo dat, zelfs als bedrijf B de belangrijkste software- of algoritme-ontwikkelaars van bedrijf A afsnoept, bedrijf A dan nog de dataset met gebruikersinformatie behoudt, waardoor A's lagere innovatiekosten blijven bestaan.

De robuustheid van een machtspositie in een datagedreven markt verschilt ook van de dominantie van Microsoft in de markt voor desktopbesturingssystemen, die gebaseerd is op directe netwerkeffecten. Directe netwerkeffecten zijn een krachtige bron van dominantie, maar kunnen worden overwonnen als de concurrent een beter product heeft, zoals Myspace heeft moeten leren van Facebook. In een datagedreven markt met indirecte netwerkeffecten kan een dominant bedrijf echter een radicaal vernieuwend product van een nieuwkomer nabootsen, tegen veel lagere innovatiekosten vanwege zijn uitgebreide en exclusieve kennis van gebruikersvoorkeuren en -kenmerken.

Zowel mededingingshandhaving als regulering kunnen de tendens van monopolisering in datagedreven markten tegengaan. Uitgangspunt daarbij is dat regulering een aanvulling is op het mededingingsrecht en nieuwe regelgevende maatregelen alleen moeten worden overwogen als vastgestelde problemen niet voldoende met het mededingingsrecht op te lossen zijn.

MEDEDINGINGSRECHT

Data vormen een niet-rivaliserend goed: als één entiteit de data heeft gebruikt, belet dit anderen niet om diezelfde informatie te gebruiken. Bovendien komt de waarde van de data vaak niet voort uit de verzamelde gegevens zelf, maar uit de kennis die eraan kan worden ontleend. Dit impliceert dat verschillende entiteiten dezelfde kennis kunnen opdoen door verschillende datasets te verzamelen.

Niettemin moet de claim van tegenstanders van verplichte datadeling, dat data vaak breed beschikbaar zijn, genuanceerd worden. Om een bepaalde marktactiviteit te beginnen, kan een specifiek type data nodig zijn, dat mogelijk niet direct beschikbaar is op de markt en ook niet replicerbaar is voor een nieuwkomer. Een voorbeeld is als er specifieke zoekgegevens nodig zijn om een zoekmachine van goede kwaliteit op te zetten, waarbij die gegevens alleen kunnen worden verkregen door zelf klanten te bedienen. Data die verkregen zijn van derden zullen dan vaak geen adequaat substituuut vormen voor de zoekgegevens van de dominante speler (Graef, 2016).

Voor toepassing van het mededingingsrecht moet dan worden vastgesteld of er sprake is van misbruik van een economische machtspositie. Als dat het geval is, kunnen de mededingingsautoriteiten op de markt ingrijpen. Wanneer een onderneming met een economische machtspositie een concurrent de toegang weigert tot gebruikersinformatie die nodig is voor het opzetten van een eigen dienst, kan een dergelijke weigering in bepaalde omstandigheden neerkomen op misbruik in strijd met artikel 102 van het Verdrag betreffende de werking van de Europese Unie.

De belangrijkste voorwaarde voor het vaststellen van mededingingsaansprakelijkheid bij een weigering tot het leveren van gebruikersinformatie, is diens onontbeerlijkheid als input. Volgens het *Bronner*-arrest van het Europese Hof van Justitie vereist onontbeerlijkheid het bestaan van

technische, juridische of zelfs economische obstakels die duplicatie onmogelijk of zelfs onredelijk moeilijk maken. Het zou op zijn minst nodig zijn om vast te stellen dat het economisch niet haalbaar is om een alternatieve dienst te creëren op een schaal die vergelijkbaar is met die van de dominante onderneming.

Dit doet de vraag rijzen of data die zijn gegenereerd op markten gekenmerkt door sterke indirecte netwerkeffecten, zoals zoekmachines en sociale netwerken, eerder als onontbeerlijk kunnen worden beschouwd, omdat het op basis van de economische kenmerken van de markt voor derden vrijwel onmogelijk is om een even grote en gevarieerde dataset op te bouwen. Facebook heeft bijvoorbeeld zulke gedetailleerde informatie over gebruikers dat het deze gebruikers in meer dan 29.000(!) categorieën kan verdelen. Het is moeilijk voor te stellen dat nieuwkomers op deze markt in staat zijn effectief te concurreren (ProPublica, 2016).

REGULERING

De reikwijdte van het mededingingsrecht is dus beperkt. Toegang tot de dataset van een dominante onderneming kan alleen van geval tot geval worden opgelegd, in specifieke omstandigheden waarin een leveringsweigering neerkomt op misbruik. Regelgeving die verder gaat dan het mededingingsrecht vereist een afweging van verschillende publieke belangen, waaronder innovatie, concurrentie en gegevensbescherming.

Innovatie

Wat mededinging en innovatie betreft, moet een afweging gemaakt worden tussen de korte- en langetermijngevolgen van regelgeving. Verplichte datadeling zal op korte termijn de concurrentie en innovatie vergroten, maar kan deze op lange termijn schaden. Sommigen vrezen dat dominante bedrijven daardoor hun prikkels verliezen om in nieuwe diensten te investeren. En concurrenten worden op hun beurt niet langer gestimuleerd om alternatieve diensten te ontwikkelen als ze gegarandeerd toegang hebben tot de gegevens van het dominante bedrijf.

Verplichte datadeling blijkt positieve netto-effecten op innovatie en welvaart te hebben wanneer de datagedreven indirecte netwerkeffecten in de markt voldoende sterk zijn. PS stelden dit vast – op basis van het eerdere idee van Argenton en Prüfer (2012) – en vonden dat dit zelfs geldt in een dynamisch model waarin concurrenten weten dat hun innovatie-investeringen vandaag invloed hebben op hun marktaandeel, en dus op hun innovatiekosten van morgen. Zowel de dominante onderneming als haar concurrenten hebben sterkere innovatieprikkels onder een regime van verplichte datadeling dan onder de status quo. Als een nieuwkomer toegang krijgt tot de gebruikersinformatie van het dominante bedrijf, moet deze immers toch nog altijd de tools ontwikkelen om iets waardevols te leren van deze data en om het product af te stemmen op de gebruikersbehoeften.

Concurrentie

Hoewel veel digitale giganten als innovatieve bedrijven gezien kunnen worden, is er geen bewijs dat heviger concurrentie op datagedreven markten zal leiden tot minder

innovatie. Wel heeft empirisch onderzoek aangetoond dat er een omgekeerde U-curve bestaat (met innovatie op de y -as en concurrentie op de x -as), waar meer concurrentie alleen tot op een bepaald punt tot meer innovatie leidt. Naarmate concurrentie op de markt verder toeneemt, zal de innovatie afnemen (Aghion et al., 2005). Vanwege het geconcentreerde karakter van datagedreven markten, met relatief weinig concurrentie, bevinden deze markten zich echter aan de uiterste linkerkant van de omgekeerde U-curve, waardoor de verwachting is dat een toename van concurrentie tot meer innovatie zal leiden.

Gegevensbescherming

Tegelijkertijd moet het fundamentele recht op gegevensbescherming in het oog worden gehouden. Het delen van gebruikersinformatie heeft immers betrekking op geïdentificeerde of identificeerbare natuurlijke personen. Vooraf anonimiseren van data zou een oplossing vormen, ware het niet dat de technologie om individuen opnieuw te identificeren steeds beter wordt. Hierdoor is het moeilijk vol te houden dat informatie in datasets volledig anoniem kan zijn. Zorgen over gegevensbescherming moeten echter niet als excuus worden gebruikt om helemaal niet in te grijpen.

Het recht op gegevensportabiliteit, onlangs geïntroduceerd door de Algemene Verordening Gegevensbescherming (AVG), geeft de betrokkenen het recht om persoonsgegevens over te dragen van de ene gegevensbeheerder naar de andere in een gestandaardiseerd format. Dit nieuwe recht kan concurrentie bevorderen omdat het lock-in vermindert door gebruikers makkelijker te laten overstappen naar andere diensten. Het is bovendien goed voor te stellen dat bedrijven dit recht proactief zullen gebruiken, door mensen te bewegen om hun gegevens bij hun vorige aanbieder op te vragen (en hun daarvoor kortingen aan te bieden). Dit heeft betrekking op het idee dat men het genereren van gegevens dient te beschouwen als arbeid die moet worden betaald (Arrieta-Ibarra et al., 2018).

De AVG vereist echter niet dat persoonsgegevens verwijderd worden zodra de betrokkene een dergelijk verzoek voor gegevensportabiliteit ingediend heeft. Als gevolg daarvan kan een bedrijf met een machtspositie, zelfs als gebruikersinformatie wordt overgedragen, nog steeds toegang hebben tot meer gegevens dan zijn concurrenten. De betrokkene kan tegelijkertijd een beroep doen op het recht om vergeten te worden, maar vanwege het beperkte toepassingsgebied van het laatste recht is de afstemming tussen deze twee onafhankelijke rechten niet perfect (Graef et al., 2017).

Wel ontstaat er een positieve externaliteit wanneer gebruikers overstappen. Voor elke nieuwe gebruiker kan de aanbieder de andere gebruikers een betere dienst bieden. Omdat individuele gebruikers deze externaliteit echter niet meenemen in hun overwegingen is de verwachting dat er te weinig gebruikers zijn die overstappen. Daarom is het onwaarschijnlijk dat het recht op gegevensportabiliteit – ook al zou het veelvuldig gebruikt worden gebruikt – het kantelen van de markt tegengaat. Niettemin illustreert dit nieuwe recht, door het delen van data te bevorderen onder gegevensbeschermingsrecht, dat de belangen van innovatie en gegevensbescherming op zichzelf niet onvereenigbaar zijn.

VERPLICHTE DATADELING: EEN VOORUITBLIK

Inspiratie over hoe een verplicht datadelingsregime moet worden vormgegeven, is in specifieke sectoren op te doen. De meest leerzame parallel kan worden getrokken met de betalingsrichtlijn Payment Services Directive 2 (PSD2) in de financiële sector. Deze richtlijn geeft derde partijen, met toestemming van de rekeninghouder, toegang tot betaalrekeningen om betalingstransacties via een internettoepassing te initiëren of tot betaalrekeninginformatie om een of meerdere accounts in één applicatie samen te voegen. Als zodanig is PSD2 een perfect voorbeeld van EU-regelgeving dat een gelijk speelveld creëert. Banken moeten nu verplicht fintech-startups toegang geven tot de rekeninginformatie van de klant. Voor het implementeren van dit *Access to Account* (XS2A) heeft de Europese Bankautoriteit verschillende richtlijnen en technische reguleringsnormen opgesteld om te verduidelijken welke stappen banken moeten ondernemen, en zijn er verschillende initiatieven ontstaan om gemeenschappelijke API-standaarden te definiëren.

In andere sectoren zijn ook beleidsinitiatieven zichtbaar die als doel hebben om meer openheid te creëren. De Europese Commissie ontving naar verluidt verontrustende signalen over de toegang tot gegevens in de landbouw- en de automobielsector, waar belanghebbenden bezorgd zijn over het feit dat gegevens onder de controle blijven van fabrikanten, ten koste van de ontwikkeling van respectievelijk 'smart farming' en nieuwe toepassingen op basis van voertuiggegevens.

Al deze beleidsinitiatieven tonen de geleidelijke verschuiving naar het openstellen van markten met als doel innovatie te stimuleren, terwijl het idee van verplichte datadeling verder ontwikkeld wordt door academici, waaronder Mayer-Schönberger en Ramge (2018): "Voortbouwend op dit idee [van PS], stellen wij een progressief datadelingsmandaat voor. Dit wordt actief zodra het marktaandeel van een bedrijf een drempel – van zeg tien procent – overschrijdt. Op dat moment moet het bedrijf een willekeurig gekozen deel van zijn dataset delen met elke andere speler in de markt die erom vraagt. Hoeveel data er beschikbaar moeten worden gesteld, is afhankelijk van de omvang van het marktaandeel. Hoe dominantier een bedrijf wordt, hoe meer gegevens het zou moeten delen met de concurrentie."

Een aanvullende vraag over het organiseren van gegevensuitwisseling tussen concurrenten is welke markten nu precies onderhevig moeten zijn aan datadeling. PS toonden al aan dat op markten waar datagedreven indirecte netwerkeffecten belangrijk zijn, verplichte datadeling positieve netto-welvaartseffecten heeft, terwijl dat argument minder overtuigend is in markten waar dit effect een ondergeschikte rol speelt. Ter illustratie is het denkbaar dat

indirecte netwerkeffecten van groot belang zijn in markten voor zoekmachines en voor sociale netwerken. In retail-elektriciteitsmarkten daarentegen, waar slimme energiemeters energieleveranciers in staat stellen om exclusieve informatie te verzamelen over consumptiepatronen van kopers op huishoudniveau, kunnen voor een bedrijf andere vormen van innovatie belangrijker zijn om succesvol te zijn.

CONCLUSIE

Deze voorbeelden zijn met name gebaseerd op intuïtie en niet op solide empirisch onderzoek. Bijgevolg is de belangrijkste volgende stap in de richting van een beleid dat de negatieve gevolgen van monopolisering op datagedreven markten beperkt het uitvoeren van een reeks van empirische studies in verschillende sectoren. Daaruit moet blijken welke markten meer en welke minder datagedreven zijn, en waar dus het verplicht delen van gegevens welwaartsbevorderend kan werken zonder innovatieprikkelers aan te tasten.

LITERATUUR

- Aghion, P., N. Bloom, R. Blundell et al. (2005) Competition and innovation: an inverted-U relationship. *The Quarterly Journal of Economics*, 120(2), 701–728.
- Argenton, C. en J. Prüfer (2012) Search engine competition with network externalities. *Journal of Competition Law and Economics*, 8(1), 73–105.
- Arrieta-Ibarra, I., L. Goff, D. Jiménez-Hernández et al. (2018) Should we treat data as labor? Moving beyond 'Free'. *AEA Papers and Proceedings*, 108, 38–42.
- Camps, M. (2018) Scherp zijn bij vervagende grenzen. *ESB*, 103(4757), 6–9.
- EC (2017) Mededeling van de Commissie aan het Europees Parlement, de Raad, het Europees Economisch en Sociaal Comité en het Comité van de Regio's, 'Bouwen aan een Europese data-economie'. COM(2017) 9 final.
- Graef, I. (2016) *EU competition law, data protection and online platforms: data as essential facility*. Alphen a/d Rijn: Kluwer Law International.
- Graef, I., M. Husovec en N. Purtova (2017) *Data portability and data control: lessons for an emerging concept in EU law*. TILEC Discussion Paper, 2017-041.
- He, D., A. Kannan, R.P. McAfee et al. (2017) Scale effects in web search. *International Conference on Web and Internet Economics – WINE 2017*, 10660, 294–310.
- Junqué de Fortuny, E., D. Martens en F. Provost (2013) Predictive modeling with big data: is bigger really better? *Big Data*, 1(4), 215–226.
- Mayer-Schönberger, V. en T. Ramge (2018) *Reinventing capitalism in the age of big data*. Londen: John Murray.
- McAfee, P., J. Rao, A. Kannan et al. (2015) *Measuring scale economies in search*. Diavoorstelling te vinden op www.learconference2015.com.
- ProPublica (2016) *Facebook doesn't tell users everything it really knows about them*. Artikel te vinden op www.propublica.org.
- Prüfer, J. en C. Schottmüller (2017) *Competing with big data*. CentER Discussion Paper, 2017-007.
- The Economist (2017) The world's most valuable resource is no longer oil, but data. *The Economist*, 6 mei.

In het kort

- ▶ Datadeling heeft positieve welvaartseffecten in markten met sterke datagedreven indirecte netwerkeffecten.
- ▶ Beleidsinitiatieven vertonen een trend naar het openstellen van private datasets buiten het mededingingsrecht om.
- ▶ Meer empirisch onderzoek is nodig om te bepalen welke sectoren baat zouden hebben bij verplichte datadeling.